

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

PROCEDE DE REMPLACEMENT SPECIFIQUE D'UNE COPIE D'UN GENE PRESENT DANS LE GENOME RECEVEUR PAR L'INTEGRATION D'UN GENE DIFFERENT DE CELUI OU SE FAIT L'INTEGRATION

5

L'invention concerne un procédé de remplacement spécifique d'une copie d'un gène présent dans le génome d'un organisme eucaryote receveur par l'intégration d'un gène différent du gène inactivé.

10

De préférence, le gène receveur sera présent en au moins 2 exemplaires dans la cellule hôte transfectée. Le gène receveur est défini comme étant le gène ou se fera l'insertion du gène différent.

15

Plus particulièrement, l'invention concerne la production d'animaux transgéniques dans lesquels le gène étranger a été introduit d'une manière ciblée pour permettre, à la fois, le maintien des fonctions génétiques normales de l'animal et l'expression du gène étranger sous le contrôle de promoteurs endogènes.

20

Par "gène différent ou étranger" on entend toute séquence nucléotidique correspondant à la totalité ou à une partie d'un gène "étranger ou différent" du gène récepteur telle qu'elle est trouvée normalement dans le génome (ARN ou ADN), ou elle correspond également à une séquence modifiée artificiellement du gène normal ou encore à un fragment de cette séquence.

25

L'invention concerne également le procédé de production de ces animaux transgéniques.

30

Dans la production d'animaux transgéniques, les méthodes conventionnelles utilisées pour l'introduction de séquences d'ADN hétérologues dans la lignée cellulaire germinale, ne permettent pas de contrôler le site de l'intégration du gène étranger dans le génome, ni le nombre de copies ainsi

**FEUILLE DE REMPLACEMENT**

intrduit. L'intégration du gène étranger se fait au hasard et, en général, plusieurs copies du gène s'intègrent en même temps, parfois sus forme de tandem tête à queue, le site de l'intégration et le 5 nombre de copies intégrées variant d'un animal transgénique à l'autre.

Il peut donc arriver que des gènes cellulaires endogènes, situés au point d'insertion, soient ainsi inactivés, sans que cela soit facilement décelable en 10 raison de nombreuses insertions au hasard. Si le produit de ces gènes est important pour le développement de l'animal, celui-ci sera sérieusement perturbé. D'ailleurs, l'insertion aléatoire du gène étranger peut se faire à un site qui n'est pas 15 approprié pour l'expression du gène. De plus, le fait qu'il y ait variation du site et du nombre d'insertions d'animal en animal rend l'interprétation des études d'expression extrêmement difficile.

Un problème majeur rencontré dans la production 20 d'animaux transgéniques, est l'obtention de l'expression du gène étranger. D'une manière générale, deux types d'expérience ont été réalisés chez les souris.

Les gènes introduits dans la lignée germinale 25 sont :

- soit des gènes "complets", comprenant des séquences codantes flanquées de leurs propres séquences régulatrices;
- soit des gènes composites, formés de la séquence 30 codante d'un gène fusionnée à la séquence promotrice d'un autre gène, les deux fragments appartenant même parfois à deux espèces animales différentes.

On a pu ainsi confirmer que la spécificité de l'expression des gènes dans tel ou tel tissu est 35 déterminée par leur(s) séquence(s) régulatrices.

#### FEUILLE DE REMPLACEMENT

Le choix du promoteur approprié pour l'expression du gène étranger chez l'animal transgénique est donc d'une importance primordiale.

D'autre part, la mutagénèse dirigée de gènes 5 murins dans des cellules souches embryonnaires a récemment été réalisée en faisant appel à une technique de "ciblage génétique" (gene targeting) (Thomas et al., 1987; Thompson et al., 1989).

Dans le premier cas, le gène murin HPRT a été 10 muté par insertion et remplacement et, dans le deuxième cas, un gène HPRT muté a été corrigé. Thomson et al. ont étendu leurs expériences jusqu'à l'obtention de souris chimères et ont constaté le passage de la modification génétique dans la lignée 15 cellulaire germinale.

Dans chacun des documents cités, le site précis de l'intégration a été ciblé par recombinaison homologue entre, d'une part, des séquences exogènes comportant la mutation ou correction incluses dans un 20 vecteur, sous le contrôle d'un promoteur exogène, et, d'autre part, leur homologue génomique. Ceci étant, il faut remarquer que les auteurs antérieurs ont réalisé leurs expériences sur un gène spécifique (HPRT) dont l'activation par mutation s'accompagnait 25 d'un phénotype décelable. La mutation ciblée décrite par Thomas et al., avait pour effet d'inactiver le gène HPRT et, par conséquent, de faire disparaître le phénotype décelable normalement associé avec le HPRT. Le gène de sélection Neo<sup>R</sup>, sous le contrôle d'un 30 promoteur TK, était donc incorporé dans l'ADN d'insertion afin de permettre la sélection des transformants. Il est à noter que les expériences décrites dans l'art antérieur impliquaient une sélection soit par le gène receveur (p. ex HPRT) soit 35 par le gène d'insertion (p. ex Neo<sup>R</sup>). Le site de

#### FEUILLE DE REMPLACEMENT

l'insertion et/ou le type de gène inséré est donc limité à des gènes conférant un caractère sélectable.

En outre, dans l'art antérieur, les séquences exogènes sur le vecteur servent donc à la fois à 5 cibler le site d'intégration et à introduire la modification. Suite à la recombinaison homologue, le gène modifié se trouve toujours dans son environnement génétique normal.

Rappelons qu'un problème qui se pose au cours de 10 la production d'animaux transgéniques est le danger d'inactiver un gène cellulaire endogène qui se trouve au point d'insertion du gène étranger.

Selon la fonction du produit du gène inactivé, une telle inactivation peut conduire à des désordres 15 physiologiques ou morphologiques importants chez l'animal transgénique, ou pourrait même empêcher sa survie.

En revanche, l'inactivation d'un gène pourrait être considéré comme avantageux si le gène en 20 question codait pour un récepteur de virus ou autre agent infectieux.

Les inventeurs ont étudié la possibilité d'éviter les inconvénients décrits plus haut, et associés, dans certains cas, à l'inactivation 25 possible d'un ou plusieurs gènes cellulaires endogène de fonction importante au cours de la production d'animaux transgéniques.

L'invention a pour objet un procédé de remplacement spécifique, notamment par ciblage d'un 30 ADN, dit ADN d'insertion constitué par une partie d'un gène susceptible d'être rendu fonctionnel, ou dont le fonctionnement peut être rendu plus efficace, lorsqu'il est recombiné avec un ADN de complément pour alors fournir un gène recombinant complet dans 35 le génome d'une cellule eucaryote, caractérisé en ce que

- le site d'insertion se trouve dans un gène choisi, dit gène receveur, et contenant l'ADN de complément, et en ce que
- l'on transfec<sup>t</sup>e des cellules eucaryotes avec un vecteur contenant un insérat comprenant lui-même l'ADN d'insertion et deux séquences dites "flanquantes" de part et d'autre de l'ADN d'insertion, respectivement homologues à deux séquences génomiques qui jouxtent le site d'insertion souhaité dans le gène receveur,
- l'ADN d'insertion étant hétérologue vis-à-vis du gène receveur, et
- les séquences flanquantes, étant choisies parmi celles qui constituent le susdit ADN de complément et qui autorisent, par recombinaison homologue avec des séquences correspondantes du gène receveur, la reconstitution d'un gène recombinant complet dans le génome de la cellule eucaryote.

L'invention concerne aussi un procédé de production d'animaux transgéniques, caractérisé en ce que des cellules E.S. sont transfectées dans les conditions sus-décr<sup>i</sup>tes et sélectionnées pour l'évènement de recombinaison homologue, à savoir l'intégration correcte du gène étranger, les cellules transfectées sont injectées dans des embryons à un stade où ils sont aptes à intégrer les cellules transfectées (par exemple au stade blastocyste), ceux-ci sont ensuite réimplantés dans une mère porteuse et les individus chimères obtenus au terme de la gestation sont accouplés. Si les cellules E.S. ont colonisé la lignée germinale de l'animal chimère, des animaux transgéniques hétérozygotes pour le gène remplacé seront obtenus par accouplement (F1) dans la descendance.

Il est également possible d'insérer le gène, porté par le vecteur de l'invention, dans l'oeuf peu

FEUILLE DE REMPLACEMENT

de temps après la fécondation (c'est-à-dire moins de 24 heures). De cette manière, l'insertion est effectuée pendant que l'oeuf est à l'état unicellulaire.

5 L'invention concerne aussi un plasmide apte à effectuer l'insertion ciblée d'un gène recombinant dit gène d'insertion dans le génome d'une cellule eucaryote, caractérisé en ce qu'il contient un insérat comprenant lui-même le gène d'insertion et  
10 deux séquences dites "flanquantes" de part et d'autre du gène d'insertion, respectivement homologues aux deux séquences génomiques qui jouxtent le site d'insertion souhaité dans le gène receveur.

L'invention concerne également des animaux  
15 transgéniques dans lesquels au moins un gène endogène a été inactivé par l'insertion d'un gène qui est différent du gène inactivé, le gène d'insertion étant inséré dans une position qui permet l'expression de ce gène sous le contrôle des séquences régulatrices  
20 du gène endogène inactivé.

Le procédé de l'invention permet, donc, grâce au phénomène de recombinaison homologue, d'insérer d'une manière ciblée des gènes étrangers, en particulier des séquences codantes dépourvues du promoteur qui  
25 leur est normalement associé, dans le génome d'un organisme eucaryote à un site qui permet leur expression sous le contrôle du promoteur endogène du gène où se fait l'insertion; et par conséquence, d'inactiver le gène endogène ciblé.

30 Selon un mode de réalisation préféré de l'invention, le gène receveur ciblé est un gène qui est présent dans le génome en au moins deux exemplaires. L'utilisation de la technique d'électroporation (Ref. 11) assure l'introduction  
35 d'une copie seulement du gène étranger.

#### FEUILLE DE REMPLACEMENT

Selon cette variante de l'invention, l'insertion ciblée du gène d'intérêt (c'est-à-dire le gène dit d'insertion) a pour effet d'inactiver la seule copie du gène cellulaire endogène où se fait l'insertion et 5 laisse intacte et fonctionnelle la ou les autre(s) copie(s) de ce gène.

De cette façon, le fonctionnement génétique de l'animal transgénique n'est pas ou peu perturbé par l'introduction du gène étranger, même si l'insertion 10 inactive une seule copie d'un gène essentiel receveur pour le développement de l'animal. Soit son développement ne serait donc pas affecté par l'insertion du gène étranger, soit les perturbations mineures possibles dans le cas de l'inactivation d'un 15 gène critique ne seraient probablement pas létales pour l'animal. Les effets de l'insertion du gène étranger à l'état homozygote pourraient être de toute nature et seraient observés en 2ème génération (F2) après croisements d'individus hétérozygotes (F1) 20 entre eux.

Si, par contre, l'inactivation de toutes les copies d'un gène est souhaitée, par exemple, dans le cas où le gène code pour un récepteur d'agent infectieux, de multiple copies du gène étranger sont 25 introduites. Le contrôle de la quantité introduite peut être assuré en faisant appel à des techniques connues.

L'insertion ciblée du gène étranger permet donc son introduction dans un site où son expression est 30 sous le contrôle des séquences régulatrices du gène endogène où se fait l'insertion.

Le procédé de l'invention permet ainsi d'insérer le gène étranger derrière un promoteur endogène qui a les fonctions désirées (par exemple, spécificité 35 d'expression dans tel ou tel tissu), et cela, le cas

#### FEUILLE DE REMPLACEMENT

échéant, sans inactiver les autres copies du gène receveur.

Selon un mode de réalisation particulièrement préféré de l'invention, l'ADN d'insertion comporte entre les 5 séquences flanquantes, d'une part une séquence d'ADN destinée à être recombinée avec l'ADN de complément dans le gène receveur pour fournir un gène recombinant, et, d'autre part, une séquence codant pour un agent sélectif permettant la sélection des 10 transformants et un promoteur autorisant l'expression de l'agent sélectif, le gène receveur et le gène recombinant codant pour des produits d'expression ne conférant pas de phénotype sélectable.

De cette manière, la sélection des transformants 15 est entièrement indépendante de la nature du gène receveur et du gène inséré, contrairement aux procédés décrits jusqu'à ce jour dans lesquels le gène inséré ou le gène receveur devait par nécessité coder pour un produit d'expression permettant la 20 sélection des transformants. Le système développé par les inventeurs permet une flexibilité totale en ce qui concerne la nature du gène receveur et du gène inséré ou du gène formé par la recombinaison homologue. Les inventeurs ont constaté d'une manière 25 surprenante que l'insertion de séquences de taille importante (par exemple d'environ 7.5 kb) n'affecte pas la fréquence de recombinaison homologue.

L'effet que peut avoir l'insertion de la séquence d'ADN selon cet aspect de l'invention inclut, selon 30 le type de séquence insérée par exemple le remplacement d'une séquence codante, le remplacement d'une séquence régulatrice, l'inactivation ou la réactivation d'un gène par mutation ou l'amélioration du taux d'expression d'un gène. Il est possible, 35 selon l'invention, de remplacer une phase codante ou une partie d'une phase codante par une séquence

FEUILLE DE REMPLACEMENT

hétérologue qui commence au codon d'initiation du gène remplacé afin que l'expression du gène inséré remplace entièrement l'expression du gène remplacé. Ceci évite la formation de protéines de fusion qui 5 pourrait être indésirable chez un animal transgénique.

Selon ce mode de réalisation de l'invention, l'ADN d'insertion peut comporter entre les séquences flanquantes une séquence codante hétérologue 10 dépourvue de promoteur, la séquence codante étant autre qu'un gène codant pour un agent de sélection. L'ADN d'insertion peut comporter en outre, en aval de la séquence codante et toujours entre les séquences flanquantes, un gène codant pour un agent de 15 sélection, associé à un promoteur permettant son expression dans la cellule cible.

De cette manière, la séquence codante hétérologue peut être insérée derrière un promoteur endogène qui a les propriétés souhaitées, par exemple une certaine 20 spécificité d'expression, ou grille de transcription etc, la sélectabilité des cellules transformées étant entièrement indépendante de l'expression de la séquence codante hétérologue. Ce type de construction permet, par exemple, de sélectionner les 25 transformants même si le gène remplacé par la séquence codante hétérologue n'est pas normalement exprimé dans les cellules cibles. Ceci est particulièrement important dans la production 30 d'animaux transgéniques à partir de cellule E.S. ("Embryonic Stem Cells") puisqu'une proportion importante des gènes reste inactive jusqu'à un stade plus avancé du développement de l'animal. Le gène Hox-3.1 est un exemple de ce type de gène. D'autre part, si la séquence codante code pour une protéine 35 facilement décelable, par exemple, le  $\beta$ -Gal, le développement de la grille de transcription du gène

#### FEUILLE DE REMPLACEMENT

endogène remplacé peut être suivi. Le vecteur pGN est un exemple de ce type de construction.

Selon un autre mode de réalisation de l'invention, l'ADN d'insertion peut comporter une 5 séquence régulatrice étrangère. Le site d'insertion et, par conséquence, les séquences flanquantes sont choisies en fonction du but désiré, à savoir soit l'insertion de la séquence régulatrice étrangère pour donner un effet de "double promoteur" avec la 10 séquence régulatrice endogène, ou soit le remplacement d'un promoteur endogène par le promoteur étranger. La séquence codante qui se trouve sous le contrôle de la séquence régulatrice peut être endogène.

15 Une autre possibilité serait l'insertion ciblée d'un ADN étranger qui comporte à la fois une séquence régulatrice et une séquence codante. Il est possible que la séquence régulatrice soit celle qui est naturellement associée avec la séquence codante.

20 Le procédé de l'invention met en œuvre un vecteur contenant deux séquences "flanquantes" de part et d'autre du gène étranger. Ces séquences flanquantes ont au moins 150 paires de bases et sont de préférence inférieures à la longueur du gène 25 receveur. Il est essentiel que ces deux séquences flanquantes soient homologues aux deux séquences génomiques qui jouxtent le site d'insertion souhaité. La séquence flanquante du vecteur qui se trouve en amont du gène étranger à introduire, est normalement 30 homologue à la séquence génomique qui est située du côté 5' du site d'insertion. De la même manière, la séquence flanquante du vecteur qui se trouve en aval du gène étranger, est normalement homologue à la 35 séquence génomique qui est située au côté 3' du site d'insertion.

#### FEUILLE DE REMPLACEMENT

11

Il est possible d'introduire des séquences "intercalantes" entre l'une ou l'autre des séquences flanquantes et le gène étranger, par exemple des séquences permettant la sélection des transformants, 5 des marqueurs, des séquences permettant le clonage du vecteur, etc...

La position de ces séquences intercalantes vis-à-vis du gène étranger doit pourtant être choisie afin de ne pas empêcher l'expression du gène 10 étranger, en particulier de la séquence d'ADN codante étrangère sous le contrôle du promoteur endogène ou, à l'inverse, la séquence codante d'ADN endogène sous le contrôle d'éléments de régulation étrangers apportés par la séquence d'insertion.

15 Malgré la présence des séquences flanquantes, qui encourage une recombinaison homologue, il est possible qu'un certain nombre d'intégrations se fasse au hasard. Afin de vérifier que l'insertion ciblée a bien eu lieu dans le site ciblé et non pas dans un autre endroit, la technique du "Polymerase Chain Reaction" (P.C.R.) (voir Ref. 10) est utilisée pour 20 amplifier la séquence d'ADN du locus où l'insertion aurait dû se faire. De cette façon, seuls les clones transformés à la suite d'une recombinaison homologue 25 sont sélectionnés.

Les séquences flanquantes du vecteur sont bien évidemment choisies en fonction du site d'insertion désiré pour que la recombinaison homologue puisse avoir lieu. Le cas échéant, les séquences flanquantes 30 peuvent comporter des séquences répliques du promoteur endogène et/ou des modifications aux séquences qui précèdent le codon d'initiation pour améliorer le taux de traduction (séquences amont) et 35 séquences répliques des séquences de terminaison notamment des sites de polyadénylation (séquences aval).

FEUILLE DE REMPLACEMENT

Le gène d'insertion peut être n'importe quel gène d'intérêt. On citera comme exemples non-limitatifs, le gène lac.Z (comme dans le modèle décrit plus loin), les gènes codant pour l'interleukine ou l'interféron, les gènes de récepteur, par exemple de l'acide rétinoïque ou beta-3 adrénnergique ou de H.I.V., et des gènes connus comme étant liés à certaines maladies, par exemple la myopathie, etc...

10 Selon une variante préférée de l'invention, les cellules eucaryotes sont des cellules souches embryonnaires (voir Ref. 14 et 15).

En effet, une cellule E.S. mutée peut être injectée dans un embryon précoce qui, après 15 réimplantation, pourra naître sous une forme chimère. Si la lignée germinale est colonisée par la cellule mutée, l'animal chimère transmettra la mutation à sa descendance. Par la suite, on pourra observer les effets de cette mutation, à l'état homozygote chez 20 certains individus, sur leur développement, leur comportement, leur métabolisme, leur pathologie, etc..

La figure 1 montre le plasmide pGN.

Les figures 2 a et b montrent les molécules pGMA 25 et pGMD respectivement construites à partir du plasmide pGN par rapport au gène Hox-3.1. Ces plasmides sont des plasmides de mutagénèse. Les deux parties de la phase codante du gène Hox-3.1 sont représentées, sur le chromosome 15, avec la boîte "homéo" en noir. Les séquences correspondantes de 30 Hox-3.1 ont été clonées dans le plasmide pGN.

(A : signal de polyadénylation; Enh/Pro : enhancer-promoteur).

07 et 08 figurent les deux oligonucléotides utilisés 35 dans la PCR.

#### FEUILLE DE REMPLACEMENT

Les figures 3 à 6 montrent les plasmides utilisés dans la construction du pGN.

La figure 7 illustre la détection de recombinaison homologue avec la technique Réaction de Polymérase en Chaîne (P.C.R.) sur des cellules E.S. transfectées.

La figure 8 (a) et (b) montre des analyses de Southern de clones individus positifs (L5 et F2) et cellules E.S. (C.C.E.).

10 Le procédé de l'invention est d'application industrielle très large et peut varier selon la nature du gène étranger introduit.

La génétique des mammifères va progresser de manière considérable grâce à la possibilité récente de mutagéniser spécifiquement n'importe quel gène, permettant ainsi de mieux définir son rôle. Par cette technologie qui fait intervenir recombinaisons homologues et cellules E.S., des informations précieuses seront apportées sur des oncogènes, des facteurs de croissance, des facteurs de transcription, etc..., gènes qui concernent des sujets très actuels de la recherche fondamentale ou la recherche appliquée. Un débouché important pour la recherche médicale est la possibilité de reproduire une maladie humaine dont la détermination génétique est connue (certaines maladies humaines à pathologie, telle la myopathie de Duchesne) ceci afin de mieux en étudier les mécanismes et de rechercher une thérapeutique.

30 En appliquant le procédé de l'invention, un gène connu comme étant responsable d'une certaine maladie est inséré d'une manière ciblée dans le génome d'une cellule E.S. L'animal transgénique qui est produit à la suite présente un modèle utile de cette maladie.

35 Si nécessaire, et comme décrit plus haut, les fonctionnements génétiques normaux peuvent être

#### FEUILLE DE REMPLACEMENT

sensiblement maintenus, malgré l'insertion du gène étranger.

Une autre application du procédé de l'invention consiste à insérer un gène d'insertion qui est facilement décélé e.g. le gène lac.Z et qui peut donc jouer le rôle d'un marqueur cellulaire. De cette manière, des études de filiation e.g. chez des animaux de concours sont facilitées, et la race peut être suivie.

10 L'insertion du gène lac.Z comme gène d'insertion rend aussi possible des études de promoteur. Grâce à la possibilité de déceler l'activité  $\beta$ -galactosidase, l'activité et la spécificité de différents promoteurs endogènes peuvent être étudiées en ciblant différents sites dans le même type ou différents types de cellules. Les mêmes études pourront être effectuées sur un organisme entier, au cours du développement ou à l'état adulte, en utilisant les techniques d'animaux chimères ou transgéniques.

15 20 Les inventeurs ont constaté d'une manière surprenante que la fréquence de recombinaison homologue n'est pas affectée par l'insertion de fragments de taille importante, par exemple le Lac Z. Cette observation a suggéré aux inventeurs que la technique de recombinaison homologue serait bien adaptée à l'insertion d'autres gènes hétérologues qui sont de taille importante.

25 30 Grâce à la possibilité de modifier le génome d'un animal, le procédé de l'invention peut également être utilisée en tant que "thérapie génique". Les utilisations les plus évidentes consisterait à inactiver les gènes de récepteurs d'agents infectieux (virus ou bactéries) ou toxiques. Si cette mutagénèse s'avérait létale, il faudrait rétablir la fonction perdue sans rétablir la sensibilité aux agents nuisibles. Un gène modifié codant pour un tel

FEUILLE DE REMPLACEMENT

récepteur pourrait être réintroduit dans la cellule mutée, à moins que la modification puisse être provoquée par la recombinaison homologue. Cette modification du patrimoine génétique conférerait à l'animal une immunité contre la maladie considérée.

Ce protocole peut aussi intervenir dans le cadre d'auto-greffe. Des cellules, malades ou saines, prélevées sur un patient, pourraient être soignées et immunisées, puis réimplantées chez le même individu.

La technique de l'invention se prête aussi aux études d'activité de produits pharmaceutiques présumés avoir une activité à l'égard des produits d'expression d'un gène pathologique lié à une maladie. Dans ce cas, le gène d'insertion est constitué par le gène pathologique et on administre à l'animal transgénique le produit pharmaceutique en vue d'évaluer son activité sur la maladie.

L'invention va être illustrée en faisant référence au plasmide pGN et son utilisation dans l'insertion ciblée d'un gène étranger (lac.Z, codant pour l'enzyme  $\beta$ -galactosidase d'E.Coli) dans le génome d'une cellule d'E.S. de souris. Le gène lac.Z a été choisi en raison du fait que son expression peut être facilement décelée et est simplement à titre illustratif.

La phase codante de l'enzyme  $\beta$ -galactosidase d'E.Coli (lac.Z; 1-3057), fusionnée avec une séquence génomique (7292-3) du gène murin Hox.3-1 (Ref. 1), débute par le codon d'initiation de ce gène. En effet, la séquence qui précède le codon d'initiation de Hox-3.1 est identique à la séquence consensus observée chez les vertébrés (Ref. 2), permettant ainsi un meilleur taux de traduction de la  $\beta$ -galactosidase dans les cellules de vertébrés. Le gène lac.Z est suivi d'un signal de polyadénylation par

FEUILLE DE REMPLACEMENT

exemple du virus SV 40, comme la plupart des gènes eucaryotes, afin de stabiliser les ARN, messagers.

L'activité de la  $\beta$ -galactosidase d'E.C li, qui est fonctionnelle dans les cellules eucaryotes, peut être décelée de différentes manières. Des cellules exprimant le gène lac.Z prennent une coloration bleue, après fixation, en présence de X-Gal, qui est un substrat de la  $\beta$ -galactosidase (Ref. 3). Un nouveau substrat, le FDG (flurorosceine di- $\beta$ -galactopyranoside), permet de déceler et de doser l'activité  $\beta$ -gal. tout en gardant les cellules vivantes (Ref. 4). Les cellules exprimant lac.Z accumulent un produit fluorescent et peuvent être isolées à l'aide d'un trieur de cellule ou FACS (fluorescence-activated cell sorter).

L'unité de transcription du gène de résistance à la néomycine provient, en grande partie, du plasmide pRSV neo (Ref. 5). Le LTR (long terminal repeat) du virus du sarcome de Rous procure des séquences activatrice et promotrice très puissantes dans de nombreuses cellules eucaryotes (Ref. 6). Du transposon bactérien Tn5, viennent un promoteur actif dans E.Coli et la phase codante de l'enzyme phosphotransférase (Ref. 7), qui est suivie du signal de polyadénylation du virus SV40. Le même gène sous le double contrôle des promoteurs RSV et Tn5 peut conférer la résistance à la néomycine ou la kanamycine aux bactéries et la résistance au G418 aux cellules eucaryotes.

Par l'effet d'une simple mutation ponctuelle, l'unité B des séquences activatrices (enhancer) de la souche PyEC F9.1 du virus du Polyome est devenue beaucoup plus active dans différents types de cellules, et en particulier dans les cellules de carcinome embryonnaire (EC) (Ref. 8). Deux copies de cet enhancer Py F9.1 ont été insérées en tandem dans

#### FEUILLE DE REMPLACEMENT

le plasmide pGN, en amont du LTR-RSV, et dans l'orientation "promoteur tardif" de la région régulatrice du Polyome.

Afin d'améliorer le taux de traduction de la phosphotransférase, la séquence précédent le codon d'initiation a été modifiée lors d'une mutagénèse par oligonucléotide. Ainsi la séquence T T C G C A U G est devenue G C A C C A U G, correspondant beaucoup mieux à la séquence consensus d'initiation de la traduction chez les vertébrés (Ref. 2).

Les améliorations apportées à l'unité de transcription du gène de résistance à la néomycine ont pu être estimées en transfectant des cellules souches embryonnaires (ES) de souris. A molarité égale en plasmide, une construction avec les enhancer Py F9.1 a produit 7,5x plus de clones résistants au G418 que le pRSV neo et 2 à 3x plus que le pMC1 Neo décrit par Capecchi et al (réf. 13). A nouveau, le nombre de clones a été augmenté 60x, soit 450x par rapport au pRSV neo, en modifiant la séquence d'initiation de la traduction. La recombinaison homologue peut être un événement assez rare, selon les conditions expérimentales appliquées (p. ex 1/1000 pour HPRT, réf. 13). Un vecteur présentant une efficacité de sélection élevée est donc très utile, d'autant plus que les conditions d'électroporation donnent lieu principalement à l'intégration d'une seule copie.

Le plasmide pGN contient, en outre, une origine de réplication bactérienne de type colE1, pBR322, qui permet les clonages et les préparations dans E.Coli.

Enfin, un site de clonage multiple (M.C.S.), synthétisé in vitro, qui ne contient que des sites de coupure uniques dans pGN, a été inséré en amont de lac.Z, afin de faciliter les utilisations de ce plasmide.

#### FEUILLE DE REMPLACEMENT

Les séquences "flanquantes" plasmidiques qui provoquent la recombinaison homologue sont ajoutées aux extrémités du plasmide pGN après linéarisation du plasmide en amont de lac.Z, par un site du MCS (voir fig. 2). En l'occurrence, les séquences flanquantes choisies sont homologues des séquences chromosomates issues de Hox-3.1 devant ultérieurement intervenir dans la recombinaison homologue.

La figure 2 situe la molécule construite à partir du plasmide pGN par rapport au gène Hox-3.1. Dans ce cas, une recombinaison entre les séquences plasmidiques et chromosomales de Hox-3.1 resulterait en une insertion au début de la phase codante de ce gène, donc à son inactivation totale.

Le plasmide pGN rassemble plusieurs avantages pour cette méthodologie, qui est applicable à n'importe quel gène. L'évènement de recombinaison homologue pouvant être assez rare (de l'ordre de 1 pour 1000 intégrations non-homologues), il est nécessaire de pouvoir analyser un grand nombre de clones dont la résistance au G418 soit suffisamment forte pour s'exprimer dans n'importe quelle partie du génome. Les modifications apportées à l'unité de transcription de la phosphotransférase répondent parfaitement à ces problèmes. La méthode de mutagénèse par recombinaison homologue équivaut à inactiver un gène par une insertion ou une substitution, mais le plasmide pGN présente l'avantage supplémentaire de pouvoir substituer l'expression de la  $\beta$ -galactosidase à celle du gène muté. Enfin, le MCS facilite les clonages de fragments génomiques.

EXEMPLES :

I - Construction du plasmide pGN

Les plasmides intermédiaires sont numérotés selon leur étape.

FEUILLE DE REMPLACEMENT

1<sup>er</sup> étape :

Insertion d'un site Xho I dans le site Bgl I de pRSV n o

5 Insertion d'un linker Xh I dans le site Bgl I de pRSV neo, rempli par le fragment Klenow de l'ADN polymérase d'E.Coli.

2<sup>nd</sup> étape :

Insertion d'un site Cla I dans le site Nde I du plasmide p1

10 Insertion d'un linker Cla I dans le site Nde I de p1, rempli par la polymérase Klenow.

3<sup>rd</sup> étape :

Insertion d'enhancer Py F9.1 dans le site Cla I du plasmide p2

15 Insertion d'enhancer Py F9.1 Pvu II-Pvu II isolé par un site unique, Acc I, dans le site Cla I de p2. Sélection d'un clone contenant deux enhancers orienté dans le sens "promoteur tardif".

4<sup>th</sup> étape :

20 Déletion Sma I-Hpa I du plasmide p3

Les deux enzymes, donnant des extrémités "bouts-francs", peuvent être liqués directement. Cette délétion enlève l'intron de l'antigène t de SV 40, qui n'est pas très utile, et diminue de manière appréciable la taille de l'unité de transcription de la phosphotransférase.

5<sup>th</sup> étape :

Insertion d'un site Xho I dans le site Bam HI de pCH110

30 Insertion d'un linker Xho I dans le site Bam HI du plasmide pCH 110 (Pharmacia), rempli par la polymérase Klenow.

6<sup>th</sup> étape :

Insertion du 3' lac.Z-polyA SV 40 dans le plasmide p4

FEUILLE DE REMPLACEMENT

La partie 3' de la phase codante de la  $\beta$ -galactosidase, suivie du signal de polyadénylation du virus SV 40 est isolée du plasmide p5 par les sites Xho I-Aat II et clonée dans le plasmide p4 par les mêmes sites.

5 7<sup>e</sup> étape :

Insertion du 5' lac.Z dans le vecteur KS-

La partie 5' de la phase codante de la  $\beta$ -galactosidase est isolée du plasmide pMC 1871 (Pharmacia) par les sites Pst I-Sac I et clonée dans 10 le vecteur KS- (Stratagene) par les mêmes sites.

8<sup>e</sup> étape :

Fusion d'une séquence génomique Hox-3.1 avec le 5' lac.Z

Une séquence génomique du gène Hox-3.1, clonée dans 15 le vecteur KS-, est purifiée par digestions successives par l'enzyme Sac I, puis par la nucléase Mung bean et enfin par l'enzyme Apa I. Cet insert est fusionné avec la partie 5' de la phase codante de la  $\beta$ -galactosidase par clonage dans le plasmide p7 20 digéré par Apa I-Sma I. La protéine ainsi fusionnée contient le codon d'initiation de la traduction du gène Hox-3.1 suivi de la phase codante de la  $\beta$ -galactosidase (vérifiée ensuite par séquençage).

25

Met Ser Ser <b>CCAGC ATG AGC TCC</b> <b>GGTCG TAC TCG AGG</b> ↓ Sac I <b>CCAGC ATG AGC T</b> <b>GGTCG TAC</b> ↓ nucléase Mung bean <b>CCAGC ATG</b> <b>GGTCG TAC</b>	Le Pro Gly Asp Pro ATT <b>CCC GGG GAT CCC</b> <b>TAA GGG CCC CTA GGG</b> ↓ Sma I GGG GAT CCC CCC CTG GGG Met Gly Asp Pro <b>CCAGC ATG GGG GAT CCC</b> <b>GGTCG TAC CCC CTA GGG</b>
--	--

FEUILLE DE REMPLACEMENT

9<sup>e</sup> étape :

Insertion de Hox-3.1-5' lac.Z dans le plasmide p6

La fusion Hox-3.1-5' lac.Z est isolée du plasmide p8  
5 par les sites Apa I-Sac I et clonée dans le plasmide p6 par les mêmes sites. Ce clonage a pour effet de reconstituer la phase codante de la  $\beta$ -galactosidase dans sa totalité.

10<sup>e</sup> étape :

10 Insertion du gène Neo<sup>R</sup> dans le vecteur KS+

Le gène de résistance à la néomycine (promoteur bactérien et phase codante de la phosphotransférase) est isolée du pRSV neo par les sites Hind III-Eco RI et clonée dans le vecteur KS+ (Stratagene).

11<sup>e</sup> étape :

Mutagénèse de la séquence d'initiation de Neo<sup>R</sup> dans p10

La séquence d'initiation de la traduction de la phosphotransférase est modifiée pour être identique 20 à la séquence consensus observée chez les vertébrés et permettre ainsi un taux supérieur d'initiation de la traduction donc une résistance accrue au G418 pour les cellules de mammifères. La modification crée également un site Apa LI qui permet de contrôler 25 l'efficacité de la mutagénèse.



30

Un oligonucléotide  
(CTTGGTCAATCATGGTGCACGATCCTCA) comportant une région de misappariement avec la séquence du pRSV neo (soulignée) est synthétisé (Gene Assembler, 35 Pharmacia) puis phosphorylée par la polynucléotide kinase du bactériophage T4. Une matrice simple brin du plasmide p10 est préparée grâce à l'origine f1 du

FEUILLE DE REMPLACEMENT

5 plasmide KS+ et hybridée avec l'oligonucléotide de mutagénèse. Le deuxième brin est synthétisé et réparé par la polymérase Klenow et l'ADN ligase du bactériophage T4. Après transformation de bactéries, les clones mutés sont ciblés à l'aide de l'oligonucléotide marqué au  $^{32}\text{P}$ . La mutagénèse a été vérifiée en digérant par Apa LI ainsi que par séquençage.

10 12<sup>e</sup> étape :

Remplacement de la séquence d'initiation dans le plasmide p9

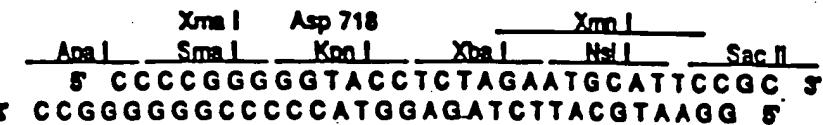
15 Un fragment contenant la séquence modifiée d'initiation de la traduction du gène de résistance à la néomycine est isolée du plasmide p11 par les enzymes Hind III-Eag I et clonée dans le plasmide p9 par les mêmes sites.

20 13<sup>e</sup> étape :

Insertion du site de clonage multiple dans le plasmide p12

25 Deux oligonucléotides complémentaires sont synthétisés (Gene Assembler, Pharmacia) puis phosphorylés. Après appariement, le MCS est cloné dans les sites Apa I- Sac II du plasmide p12 grâce à ses extrémités cohésives

30



35 Le site de clonage multiple a également été vérifié par séquençage.

II - Addition des séquences "flanquantes" aux extrémités du plasmide pGN linéarisé en amont de lac.Z' par un site du M.C.S

35 Les séquences flanquantes utilisées ont été choisies en fonction du site d'insertion souhaité

FEUILLE DE REMPLACEMENT

(par exemple, Hox-3.1, voir Fig. 2 a et b pGMA et pGMD).

Dans la construction du plasmide de mutagénèse pGMD, deux bras d'ADN homologue au locus Hox-3.1 ont été clonés aux sites Apa I-Nsi I et Nsi I-Sac II du vecteur pGN. Le bras 5' commence au site Sac II (CCGC<sub>n</sub>GG) au nucléotide 219 de l'ADNC c21 de Hox-3.1. Ce fragment s'étend sur 6.8 kb en 5' jusqu'au premier site BamHI. Le bras 3' commence au site Apa I (GGGCC<sub>n</sub>CC) au nucléotide 885 de l'ADNC c21. Ce fragment s'étend sur 1.5 kb en 3' jusqu'au premier site PstI. Un linker NsiI a été inséré dans le site BamHI du fragment 5' et dans le site PstI du fragment 3'. Les bras 5' et 3' ont été clonés dans le vecteur pGN dans les sites Nsi I-Sac II et Apa I-Nsi I, respectivement. La séquence de l'ADNC de Hox-3.1 c21 est publiée (réf. 1).

Le plasmide de mutagénèse est linéarisé par digestion avec Nsi I avant électroporation de cellules E.S. Ses extrémités sont formées des deux bras génomiques clonés aux sites Apa I-Nsi I et Nsi I-Sac II du vecteur pGN.

Le plasmide pGMD ne présente pas de signal de polyadénylation après le gène de résistance mais, en revanche, présente une séquence riche en AU responsable d'une dégradation sélective de mRNA, insérée dans la séquence de l'intron du Hox-3.1 du plasmide.

Un autre plasmide de mutagénèse, pGMA, présente la même structure que le pGMD mais contient les signaux de polyadénylation et de terminaison de transcription du SV40 et ne présente pas la séquence AU de dégradation de mRNA en aval du gène Neo'. Ces modifications avaient pour but de réduire le taux de transcrits de Neo' dans des clones issus d'intégration au hasard. En revanche, des clones

#### FEUILLE DE REMPLACEMENT

issus d'événements de recombinaison homologue entre pGMD et un locus Hox-3.1 devrait avoir une croissance inaltérée pendant la sélection au G418, la séquence AT de dégradation de mRNA étant éliminée par le 5 procédé de recombinaison même, ou épissée avec l'intron Hox-3.1.

Dans les étapes expérimentales qui suivent, le protocole décrit par Thompson et al. 1989, a été suivi pour la production d'animaux chimères.

10 III - Transfection de cellules embryonnaires de souris

La méthode décrite par Thompson et al. 1989, a été utilisée pour transfacter des cellules embryonnaires de souris. L'utilisation de la 15 technique de l'électroporation assure l'introduction d'une seule copie du gène étranger (lac.Z) par cellule. Après transfection, plusieurs clones exprimant la  $\beta$ -galactosidase ont été isolés.

Les plasmides de mutagénèse pGMD et pGMA ont été 20 linéarisés et introduits par électroporation dans des cellules E.S. afin de favoriser l'insertion d'une copie seulement dans le génome (réf. 11).

Les transfections initiales ont été effectuées pour comparer l'efficacité de ciblage du Hox-3.1 des 25 plasmides pGMA et pGMD (voir tableau I).

Recombinaison homologue dans le gène Hox-3.1				
Exp.	Plasmide de mutagénèse	N° de l'ensemble analysé	Nb de clones formant l'ensemble	Nb de résultats P.C.R. positifs
I	pGMA	3	600	0(2)
II	pGMD	5	250	3(5)
III	pGMD	84	2-3	5(5)

Tableau I

## FEUILLE DE REMPLACEMENT

La lignée cellulaire E.S. "C.C.E." (réf. 16) a été maintenue d'une manière continue sur des couches nourricières fibroblastiques (réf. 17). Pour les expériences I et II,  $1.5 \times 10^7$  cellules E.S. dans 1.5 ml HeBS ont été électroporées (réf. 11) à 200 V, avec 40 mg de plasmide linéarisé, puis étalées sur quatre boîtes de cultures (diamètre 100 mm). Pour l'expérience III, le choc a été effectué dans les mêmes conditions mais un quart des cellules ont été étalées sur quatre plaques à 24 trous. Le lendemain, 250 µg ml<sup>-1</sup> G418 ont été ajoutés. Chaque transfection a donné lieu à environ 2400 clones avec pGMA et environ 1000 clones avec pGMD.

Le nombre moyen de clones de cellules E.S. résistantes au G418 dans chaque ensemble est indiqué dans le tableau I, ainsi que le nombre d'ensembles donnant un résultat positif avec la technique P.C.R. Un résultat positif signifie q'une bande de 1.6 Kb a pu être observée sur un gel d'agarose coloré de bromure d'ethidium (voir fig. 7). Le nombre d'ensembles donnant un signal positif après une analyse de Southern du mélange P.C.R. et hybridation d'une sonde spécifique qui ne contenait pas les séquences des amorces est indiqué entre parenthèses (fig. 8).

Détection de la recombinaison homologue avec la P.C.R.

P.C.R. a été effectuée sur  $10^5$  cellules d'un ensemble de 250 clones de la transfection II (voir voie D de la fig. 7). Dans les autres voies, quatre ensembles de la transfection III ont été analysés ensemble en mélangeant environ  $4 \times 5000$  cellules. Les amorces 07 et 08 utilisées dans la P.C.R. entourent la séquence 3' Hox-3.1 du plasmide de mutagénèse (fig. 2). Le fragment de 1.6 Kb recouvrant cette séquence 3' ne peut être amplifié que dans le cas d'une

FEUILLE DE REMPLACEMENT

recombinaison homologue. Les voies 2, 3 et D illustrent des résultats positifs.

L'ADN de clones E.S. a été préparé au moment de la réplique sur filtre en utilisant la méthode "boiling-proteinase K digestion boiling" (réf. 18). 40 cycles d'amplification (40 secondes à 94°C, 1 minute à 60°C, 7 minutes à 72°C) ont été effectués dans un mélange réactionnel de 100 µl, contenant 67 mM Tris-HCl (pH 8.6), 16.7 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 6.7 mM MgCl<sub>2</sub>, 10 10 mM 2-mercaptopropanoïd, 0.01 % (p/v) gélatine, 200 µM dATP, dTTP et dCTP, 100 µM dGTP, 100 µM 7-deaza dGT, 600 ng de chaque amorce (07 : AACTTCCCTCTCTGCTATT et 08 : CAGCAGAACATACAAGCTG) et 3U polymérase Taq (Perkin Elmer Cetus), couvert de 15 100 µl paraffin. La moitié du mélange de réaction a été appliquée sur un gel d'agarose 0.7 % coloré au bromure d'ethidium. Le marqueur de taille est un digest Eco RI + Hind III d'ADN lambda.

#### Analyses de Southern

20 Trois clones indépendants de cellules E.S. contenant le Hox-3.1 muté (identifié par P.C.R.) ont été isolés des ensembles positifs en utilisant des pipettes. Leur ADN a été examiné par analyse de Southern après digestion avec les enzymes de restriction indiquées 25 dans la figure 8, afin de confirmer le ciblage spécifique et de faire la distinction entre les loci recombinés et sauvages. Deux sondes différentes ont été utilisées dans l'analyse de l'extrémité 3' des loci Hox-3.1 dans les clones mutés et dans les 30 cellules E.S. non-mutées agissant comme témoin (fig. 8 c). La première sonde "a" était contenue dans les séquences Hox-3.1 du plasmide de mutagénèse et démontrait le nombre d'intégrations de vecteur et leurs liaisons physiques. Un des trois clones 35 recombinés contenait en outre une copie du plasmide intégrée au hasard (fig. 8 a, clone F2). La deuxième

#### FEUILLE DE REMPLACEMENT

snde "b" qui n'était pas contenue dans le vecteur de mutagénèse faisait la distinction entre les allèles Hox-3.1 recombinées et sauvages (fig. 8 b). Le locus recombiné Hox-3.1 présentait, avec les deux sondes, l'image d'hybridation attendue à partir des cartes de restriction du vecteur de mutagénèse et du locus intact. En outre, l'existence de deux domaines de recombinaison dans le bras 3' du vecteur a été confirmée par la présence ou l'absence de la séquence AT dans le locus Hox-3.1 recombiné (par exemple fig. 8, clone L5). L'extrémité 5' du locus Hox-3.1 a également été analysée pour l'événement de recombinaison homologue. Des enzymes de restriction ne présentant pas de sites dans la séquence Hox-3.1 5' de 6.8 Kb du vecteur de mutagénèse ont été utilisés dans la digestion des ADNs des clones recombinés. Ces ADNs ont ensuite été soumis à une électrophorèse dans un champ pulsé pour différencier les fragments de poids moléculaire élevé. Une analyse de Southern de ce gel a également indiqué les allèles recombinées correctement et les allèles Hox-3.1 sauvages, en utilisant une sonde présentant une séquence en amont du plasmide de mutagénèse.

Les analyses de Southern ont démontré q'une allèle du gène Hox-3.1 a été recombinée comme prévu. La recombinaison homologue était équivalente à un double "crossing-over" entre les bras génomiques du plasmide de mutagénèse et les séquences chromosomales homologues (Fig. 2).

Dans les clones recombinants, le gène lac Z a été placé sous le contrôle des séquences promotrices et régulatrices du Hox-3.1 en amont du codon AUG, mais les signaux 3' de maturation de mRNA provenaient du SV40. Dans ces clones recombinés, l'expression de lac Z n'était pas décelable par coloration avec  $\beta$ -Gal, ce qui est cohérent avec l'absence de

#### FEUILLE DE REMPLACEMENT

transcription de Hox-3.1 dans des cellules E.S. déterminée par analyse de protection de RNase. L'activité de  $\beta$ -Gal pouvait être induite dans certaines cellules après 3 ou 4 jours de culture en 5 présence de  $5 \cdot 10^{-7}$  M acide rétinoïque, conditions connues comme induisant la transcription de Hox-3.1 (réf. 19).

En utilisant le vecteur de mutagénèse pGMA, qui présente une homologie totale de 8.3 Kb ADN avec le 10 locus Hox-3.1, un fragment de 120 pb a été remplacé par une insertion de 7.2 Kb. La fréquence de ce remplacement ciblé (1/900) est comparable à celle obtenue récemment (1/1000) avec HPRT (réf. 13) ou avec En-2 (1/260) (réf. 20), le fragment hétérologue 15 inséré étant cependant dans ces derniers cas d'une taille beaucoup moins importante (1.1 et 1.5 Kb respectivement). D'une manière surprenante, il a été constaté qu'une fréquence de recombinaison homologue très élevée (1/40) a pu être obtenue avec le vecteur 20 pGMD. L'élimination des signaux 3' de maturation de mRNA et l'addition de la séquence de dégradation de mRNA au gène de résistance à la néomycine a eu pour effet de réduire le nombre total de clones résistants au G418 par 2.4 (tableau I). Le rapport de ciblage 25 spécifique était presque 10 fois plus élevé (900/40). Le mécanisme de recombinaison homologue même a dû être affecté dans les expériences avec pGMD. Une explication possible de ces résultats serait qu'une séquence AT de 51 pb pourrait fournir, in vivo, une 30 boucle ouverte dans le plasmide de mutagénèse à cause de sa température de fusion plus basse. Si les séquences Hox-3.1 voisines du pGMD peuvent être influencées par cette ouverture, de chaque côté de la région AT, elles pourraient réagir d'une manière plus 35 efficace, à l'état simple-brin, avec le locus chromosomal Hox-3.1. Le modèle de recombinaison

FEUILLE DE REMPLACEMENT

mitotique chez la levure suggère qu'il serait initié par un tel échange de brins, bien que le mécanisme de recombinaison homologue reste inconnu chez les eukaryotes plus complexes.

5 La figure 8 montre les résultats de l'analyse de Southern effectuée sur des clones individus positifs (L5 et F2) et des cellules E.S. (C.C.E.).

Les sondes utilisées n'hybrident qu'avec des séquences Hox-3.1 incluses dans le vecteur (a) ou 10 exclues du vecteur de mutagénèse (b). L'image d'hybridation du locus Hox-3.1 recombiné (triangles ouverts) se distingue clairement du locus sauvage (triangles noirs). Les étoiles indiquent les bandes d'hybridation d'une copie du plasmide qui s'est 15 intégrée au hasard. Le marqueur de taille est un digest Eco RI + Hind III d'ADN lambda.

La figure 8(c) montre les cartes de restriction des allèles Hox-3.1 recombinées (rec.) et sauvages (wt). Les parties du vecteur de mutagénèse et du 20 locus Hox-3.1 sont indiquées avec les mêmes symboles que ceux utilisés dans la figure 2. Dans ce cas, la séquence AT a été intégrée par recombinaison homologue. La flèche verticale indique l'extrémité 3' du plasmide de mutagénèse. La localisation des sondes 25 "a" et "b" utilisées dans l'analyse de Southern est également indiquée. Les abréviations utilisées dans la figure 8 sont les suivantes : B, Bam HI ; D, Dra I, E, Eco RI ; H, Hind III ; S, Sal I ; X, Xba I.

#### IV - Production d'embryons chimères

30 Une microinjection dans des blastocystes a été effectuée avec deux clones E.S. recombinants contenant un allèle Hox-3.1 intacte et un allèle recombiné, ces clones ne contenaient aucune autre copie du plasmide de mutagénèse. Les Kariotypes de 35 ces cellules étaient normaux.

FEUILLE DE REMPLACEMENT

Dix à quinze cellules mutées ont été microinjectées par blastocyste. Après réimplantation dans des mères porteuses, les embryons ont été recueillis à 9.5, 10.5 et 12.5 jours p.c. et analysés pour expression de lac Z. La grille de transcription de Hox-3.1 à ces stades avait été déterminée au préalable par analyse d'hybridation *in situ* (réf. 1). Les transcrits Hox-3.1 sont détectables pour la première fois au stade de gastrulation tardive et sont répartis dans tous les tissus de la partie arrière de l'animal. Plus tard, la répartition devient progressivement limitée dans l'espace et spécifique au tissu. Au stade de 12.5 jours p.c., la transcription est localisée dans la région cervicale du tube neural, au niveau du cœur. Au cours de l'embryogenèse, la répartition de la transcription de Hox-3.1 subit donc des modifications. Le stade 10.5 jours p.c. semble être une période de transition, la transcription ayant lieu à la fois dans les deux régions arrières et dans le tube neural cervical.

Dans des embryons chimériques de 9.5 et 10.5 jours p.c., la partie caudale au bourgeon postérieur présentait une activité  $\beta$ -Gal intense, tandis que le marqueur n'a jamais été détecté dans la région thoracique antérieure ou la tête (Fig. 9a). Dans la région arrière, des cellules colorées par le  $\beta$ -Gal ont été observées dans tous les tissus et de toutes les couches embryonnaires. Entre les deux bourgeons qui donnent les membres, des cellules colorées étaient réparties dans des zones restreintes, dans l'ectoderme superficiel (Fig. 9b), comme dans les régions arrières (Fig. 9c) et, en forme de lignes étroites ou rayures, dans le tube neural (Fig. 9b). Ces rayures présentaient une répartition irrégulière et asymétrique sur la paroi du tube neural. La transcription de Hox-3.1 n'a pas été détectée dans la

#### FEUILLE DE REMPLACEMENT

couche mince de cellules vers la fermeture du tube neural. Ces cellules n'ont peut-être pas résisté aux traitements appliqués lors de l'hybridation in situ. Il a été observé que les cellules de l'ectoderme neurale font partie, très tôt, de parties différentes du système nerveux et migrent dans une direction radiale, suivant des mouvements latéraux étroits (réf. 21). Ces résultats sont donc cohérents avec cette observation.

10 L'expression de Lac Z a donc illustré correctement la première partie de la transcription de l'homéogène Hox-3.1, c'est-à-dire dans tous les tissus des régions caudales des embryons de 9.5 et 10.5 jours p.c., et a fourni de nouvelles 15 informations concernant le mode de transcription de Hox-3.1.

En revanche, l'expression de Lac Z n'a pas été observée dans les régions cervicales du tube neural d'embryons chimères de 12.5 jours, ni dans la région 20 antérieure d'embryons de 10.5 jours ; ceci n'était pas le résultat attendu à partir des études d'hybridation in situ. La phase ultérieure de transcription de Hox-3.1 observée à partir du jour 10.5 dans les zones très localisées du tube neural 25 n'était pas mise en évidence par l'activité de  $\beta$ -Gal. Une explication possible pour ce résultat serait que, bien que l'expression de Lac Z soit sous le contrôle du promoteur Hox-3.1, les séquences 3' du Hox-3.1 sont absentes dans le gène reporteur. Il est possible 30 que des séquences 3' du codon d'initiation AUG du Hox-3.1 aient une influence sur l'expression tardive de Hox-3.1 dans le domaine antérieur. Un effet de "dosage de gène" pourrait aussi expliquer ce résultat. L'autoactivation de plusieurs homéogènes 35 chez Drosophila a été démontrée génétiquement ou

#### FEUILLE DE REMPLACEMENT

32

suggérée par la formation de complexes entre l'ADN et les protéines des homeobox.

Si le composant tardif de la transcription de Hox-3.1 dans le tube neural est maintenu par un mécanisme semblable, l'inactivation d'un allèle pourrait avoir un effet dominant dans les celules de l'ectoderme neural. Puisqu'un allèle seulement produirait la protéine Hox-3.1, le signal d'activation serait dilué sur les deux promoteurs. La réduction d'autoinactivation dans les deux loci pourrait ainsi conduire à un arrêt total de l'initiation de transcription. Ceci pourrait expliquer pourquoi aucune expression de Lac Z n'a été détectée dans la région cervicale du tube neural d'embryons de 10.5 et 12.5 jours.

V - Passage de la modification dans la lignée cellulaire germinale : production d'animaux transgéniques

Les effets en  $F_1$  et en  $F_2$  de la modification apportée par l'insertion ciblée ont été observés après reproduction des chimères. Le passage de la modification dans la lignée cellulaire germinale a été constaté.

25

30

35

FEUILLE DE REMPLACEMENT

BIBLIOGRAPHIE

1. Le Mouellic, H., Condamine, H. et Brûlet, P. (1988). Genes & Development. 2, 125-135.
2. Cavenir, D.R. (1987). Nucleic Acids Res. 15, 5 1353-1361.
3. Sanes, J.R. Rubenstein, J.L.R. et Nicolas, J.F. (1986) EMBO J. 5, 3133-3142.
4. Nolan, G.P., Fiering, S., Nicolas, J.F. et Herzenberg, L.A. (1988) Proc. Natl. Acad. Sci. USA 10 85, 2603-2607.
5. Gorman, C., Padmanabhan, R. et Howard, B.H. (1983). Science. 221, 661-553.
6. Gormann, C.M., Merlino, G.T. Willingham, M.C. Pastan, I. et Howard, B. H. (1982) Proc. Natl. 15 Acad. Sci. USA 79, 6777-6781.
7. Southern, P.J. et Berg, P. (1982) J. Mol. Appl. Genet 1, 327- 341.
8. Herbomel, P., Bourachot, B. et Yaniv, M. (1984). Cell. 39, 653-662.
9. Robertson, E.J. (1987). Teratocarcinomas and embrionic stem cells: A practical approach. IRL Press, Oxford.
10. Kahn, A. (1988). Médecine/Sciences 4, 515-518.
11. Chu, Hayakawa H., Berg. P. (1987) Nucleic Acid 25 Research 15, Nr. 3, 1311-1326.
12. Thompson, S., Clarke, A.R., Pow, A.M., Hooper, M.L., Melton, D.W., (1989) Cell, 56, 313-321.
13. Thomas, K.R., Capecchi, M.R., (1987) Cell, 51 503-512.
14. Evans, M.J., Kaufmann, M.H. (1981) Nature, 292, 30 154-155,
15. Robertson, E.J., (1986) Trends in Genetics, 9-13
16. Robertson, E., Bradley, A., Kuehn, M. & Evans, M. Nature 323, 445-448 (1986).

34

17. Robertson, E.J. in *Teratocarcinomas and Embryonic Stem Cells* (ed. Robertson, E.J.) 71-112 (IRL, Oxford, 1987).
18. Kim, H.S. & Smithies, O. *Nucleic Acids Res.* 16, 5 8887-8903 (1988).
19. Breier, G., Bucan, M., Francke, U., Colberg-Poley, A.M. & Gruss, P. *EMBO J.* 5, 2209-2215 (1986).
20. Joyner, A.L., Skarnes, W.C. & Rossant, J. *Nature* 338, 153-156 (1989).
- 10 21. McKay, R. D. G. *Cell* 58, 815-821 (1989).

15

20

25

30

35

FEUILLE DE REMPLACEMENT

## REVENDICATIONS

1. Procédé de remplacement spécifique, d'un gène notamment par ciblage d'un ADN, dit ADN d'insertion constitué par une partie d'un gène susceptible d'être rendu fonctionnel, ou dont le fonctionnement peut être rendu plus efficace, lorsqu'il est recombiné avec un ADN de complément pour alors fournir un gène recombinant complet dans le génome d'une cellule eucaryote, caractérisé en ce que
  - 5 - le site d'insertion se trouve dans un gène choisi, dit gène receveur choisi, et contenant l'ADN de complément, et en ce que
    - l'on transfecte des cellules eucaryotes avec un vecteur contenant un insérat comprenant lui-même l'ADN d'insertion et deux séquences dites "flanquantes" de part et d'autre de l'ADN d'insertion, respectivement homologues à deux séquences génomiques qui jouxtent le site d'insertion souhaité dans le gène receveur,
  - 10 - l'ADN d'insertion étant hétérologue vis-à-vis du gène receveur, et
    - les séquences flanquantes, étant choisies parmi celles qui constituent le susdit ADN de complément et qui autorisent, par recombinaison homologue avec des séquences correspondantes du gène receveur, la reconstitution d'un gène recombinant complet dans le génome de la cellule eucaryote.
2. Procédé selon la revendication 1, ledit ADN d'insertion contenant soit une séquence codante soit une séquence régulatrice, caractérisé en ce que
  - 15 - le site d'insertion se trouve dans un gène choisi dit gène receveur et en ce que
    - l'on transfecte des cellules eucaryotes avec un vecteur contenant un insérat comprenant lui-même l'ADN d'insertion et deux séquences dites "flanquantes" de part et d'autre de l'ADN

FEUILLE DE REMPLACEMENT

d'insertion, respectivement homologues à deux séquences génomiques qui jouxtent le site d'insertion souhaité dans le gène receveur,

5 - l'ADN d'insertion étant hétérologue vis-à-vis du gène receveur et,

- les séquences flanquantes étant choisies afin de permettre par recombinaison homologue selon le cas, soit l'expression de la séquence codante de l'ADN d'insertion entier sous le contrôle des séquences 10 régulatrices du gène receveur, soit l'expression d'une séquence codante du gène receveur sous le contrôle de séquences régulatrices de l'ADN d'insertion.

3. Procédé selon la revendication 1 ou 2, 15 caractérisé en ce que l'ADN d'insertion contient une séquence codante dépourvue d'élément de régulation, notamment d'un promoteur qui lui est propre.

4. Procédé selon l'une quelconque des revendications 1 à 3, caractérisé en ce que le gène 20 receveur est présent dans le génome de la cellule eucaryote en au moins deux exemplaires.

5. Procédé selon l'une quelconque des revendications 1 à 4, caractérisé en ce que chacune 25 des séquences flanquantes a une longueur supérieure à 150 paires de bases, et inférieure à la longueur du gène receveur.

6. Procédé selon l'une quelconque des revendications 1 à 5, caractérisé en ce que les 30 cellules eucaryotes sont des cellules souches embryonnaires (E.S.).

7. Procédé selon l'une des revendications précédentes, caractérisé en ce que le gène d'insertion est un gène hétérologue à l'espèce transfectée.

35 8. Procédé selon l'une des revendications précédentes, caractérisé en ce que le vecteur

FEUILLE DE REMPLACEMENT

contient des séquences intercalées entre le gène d'insertion et les séquences flanquantes.

9. Procédé selon la revendication 8, caractérisé en ce que les séquences intercalantes contiennent une 5 séquence codant pour un agent sélectif permettant la sélection des transformants et éventuellement un gène marqueur par exemple le LacZ.

10. Procédé selon l'une des revendications précédentes, caractérisé en ce que la transfection 10 est effectuée par électroporation.

11. Procédé selon l'une des revendications précédentes, caractérisé en ce que la technique de Polymerase Chain Reaction (P.C.R.) est utilisée pour amplifier la séquence d'ADN du locus où se fait 15 l'insertion pour vérifier que l'insertion a eu lieu dans le site souhaité.

12. Procédé selon la revendication 1, caractérisé en ce que l'ADN d'insertion comporte, entre les séquences flanquantes, d'une part une séquence d'ADN 20 destinée à être recombinée avec l'ADN de complément dans le gène receveur pour fournir un gène recombinant, et, d'autre part, une séquence codant pour un agent sélectif permettant la sélection des transformants et un promoteur autorisant l'expression 25 de l'agent sélectif, le gène receveur et le gène recombinant codant pour des produits d'expression ne conférant pas de phénotype sélectable.

13. Procédé de production d'animaux transgéniques, caractérisé en ce que des cellules 30 E.S. sont transfectées par le procédé selon l'une des revendications 1 à 12 et sélectionnées pour l'évènement de recombinaison homologue, à savoir l'intégration correcte du gène étranger, les cellules sont injectées dans des embryons à un stade où ils 35 sont aptes à intégrer les cellules transfectées, par exemple au stade blastocyste, ceux-ci sont ensuite

FEUILLE DE REMPLACEMENT

réimplantés dans une mère porteuse et, les individus chimères obtenus au terme de la gestation et chez lesquels est constaté la colonisation par les cellules E.S. de la lignée germinale, sont accouplés 5 pour obtenir des animaux transgéniques hétérozygotes pour le gène remplacé.

14. Plasmide apte à effectuer l'insertion ciblée d'un gène dit gène d'insertion dans le génome d'une cellule eucaryote, caractérisé en ce qu'il contient 10 un insérat comprenant lui-même le gène d'insertion et deux séquences dites "flanquantes" de part et d'autre du gène d'insertion, respectivement homologues aux deux séquences génomiques qui jouxtent le site d'insertion souhaité dans le gène receveur.

15 15. Plasmide selon la revendication 14, caractérisé en ce que l'insérat comprend, entre les séquences flanquantes, d'une part une séquence d'ADN destinée à être recombinée avec l'ADN de complément dans le gène receveur, et, d'autre part, une séquence 20 codant pour un agent sélectif permettant la sélection des transformants et un promoteur autorisant l'expression de l'agent sélectif, la séquence d'ADN destinée à être recombinée avec l'ADN de complément étant autre qu'un gène codant pour un agent sélectif.

25 16. Plasmide pGN comme illustré dans la figure 1.

17. Cellules eucaryotes transformées par le procédé de la revendication 1.

18. Cellules selon la revendication 17 30 caractérisées en ce que ce sont des cellules E.S.

19. Animal transgénique dans lequel une seule copie d'un gène qui est présent dans le génome à au moins deux exemplaires, a été inactivé par l'insertion d'un gène qui est différent du gène 35 inactivé, le gène d'insertion étant inséré dans une position qui permet l'expression de ce gène sous le

FEUILLE DE REMPLACEMENT

contrôle des séquences régulatrices du gène endogène inactivé.

20. Application du procédé selon l'une quelconque des revendications 1 à 12 pour la thérapie 5 génique.

21. Application du procédé selon l'une quelconque des revendications 1 à 12 pour la production d'animaux transgéniques.

22. Application du procédé de la revendication 9 10 pour marquage génétique d'animaux.

23. Application du procédé de la revendication 13 pour le criblage de produits pharmaceutiques présumés avoir une activité à l'égard des produits 15 d'expression d'un gène pathologique lié à une maladie, caractérisée en ce que le gène d'insertion est constitué par le gène pathologique ou un fragment de celui-ci et en ce que l'on administre à l'animal transgénique le produit pharmaceutique à tester, en vue d'évaluer son activité sur la maladie.

20

25

30

35

FEUILLE DE REMPLACEMENT

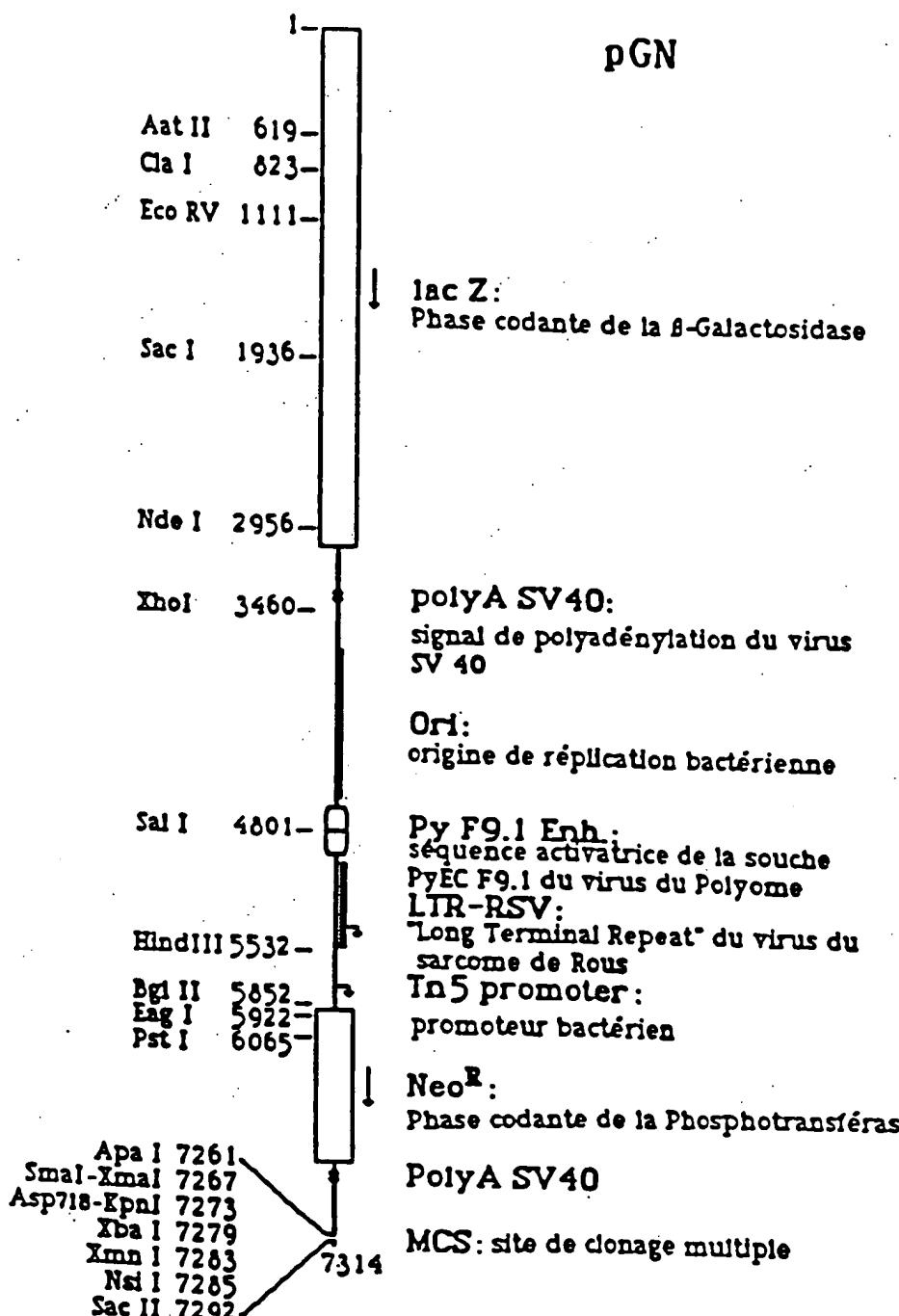


Figure 1: plasmide linéaire

EFFICACIE DU REMPLACEMENT

2A2

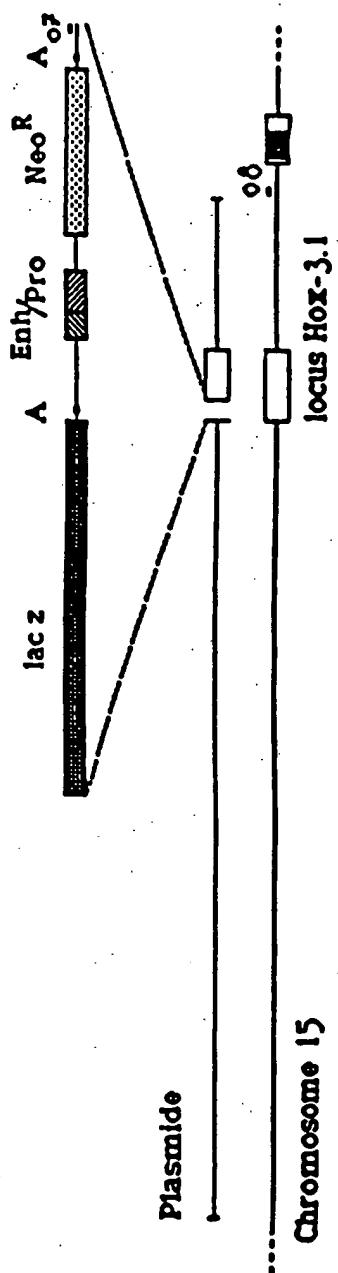


FIGURE 2a

SCHEME OF REPLACEMENT

3/12

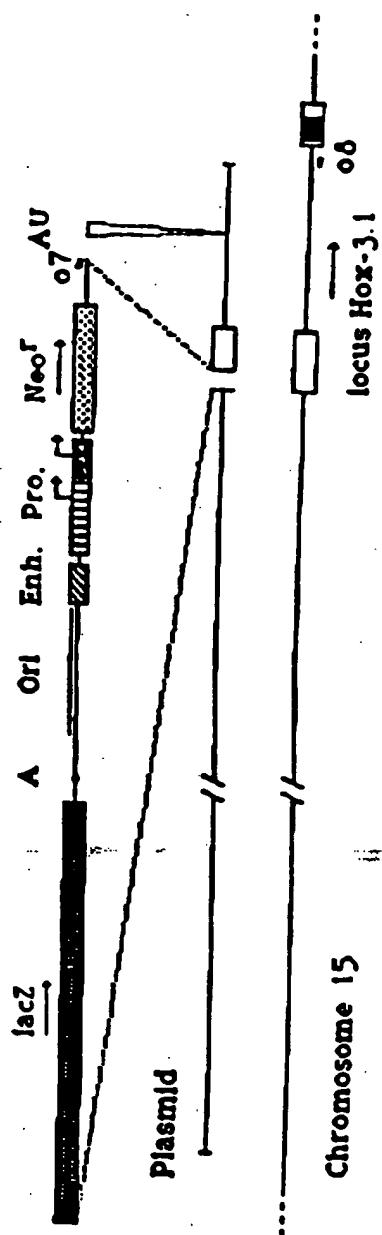


FIGURE 2b

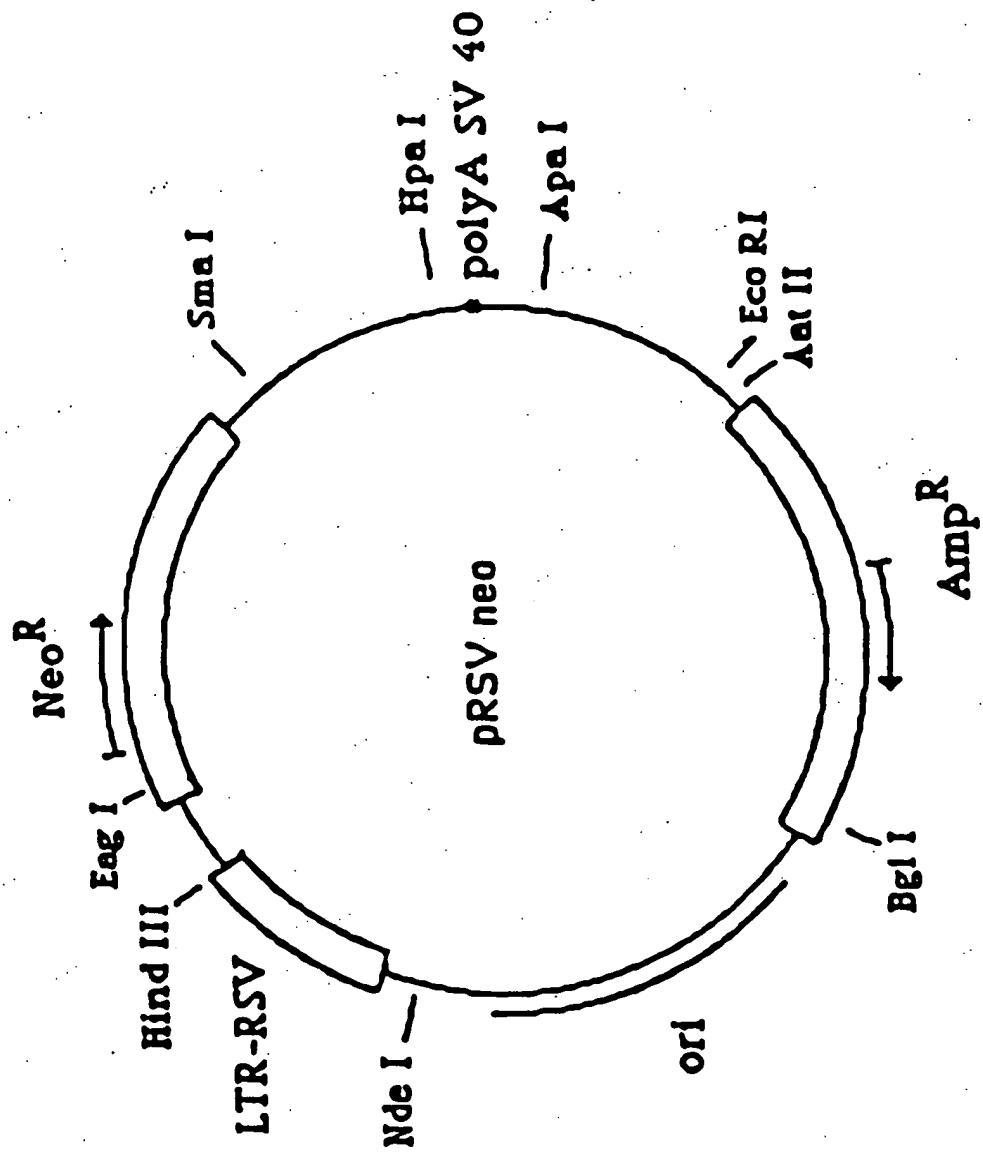
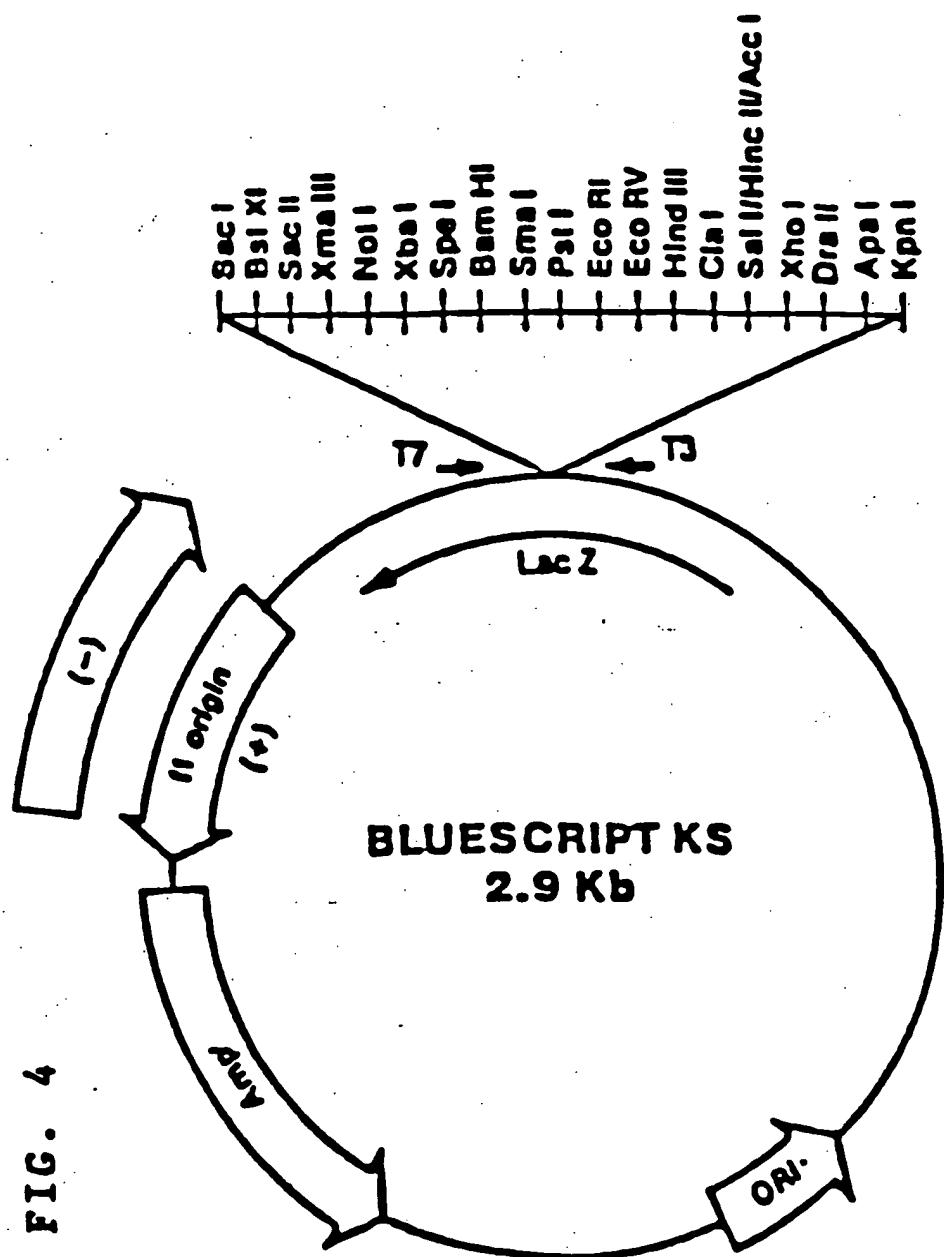


FIG. 3

EXAMPLE OF REPLACEMENT



FEUILLE DE REMPLACEMENT

6/12

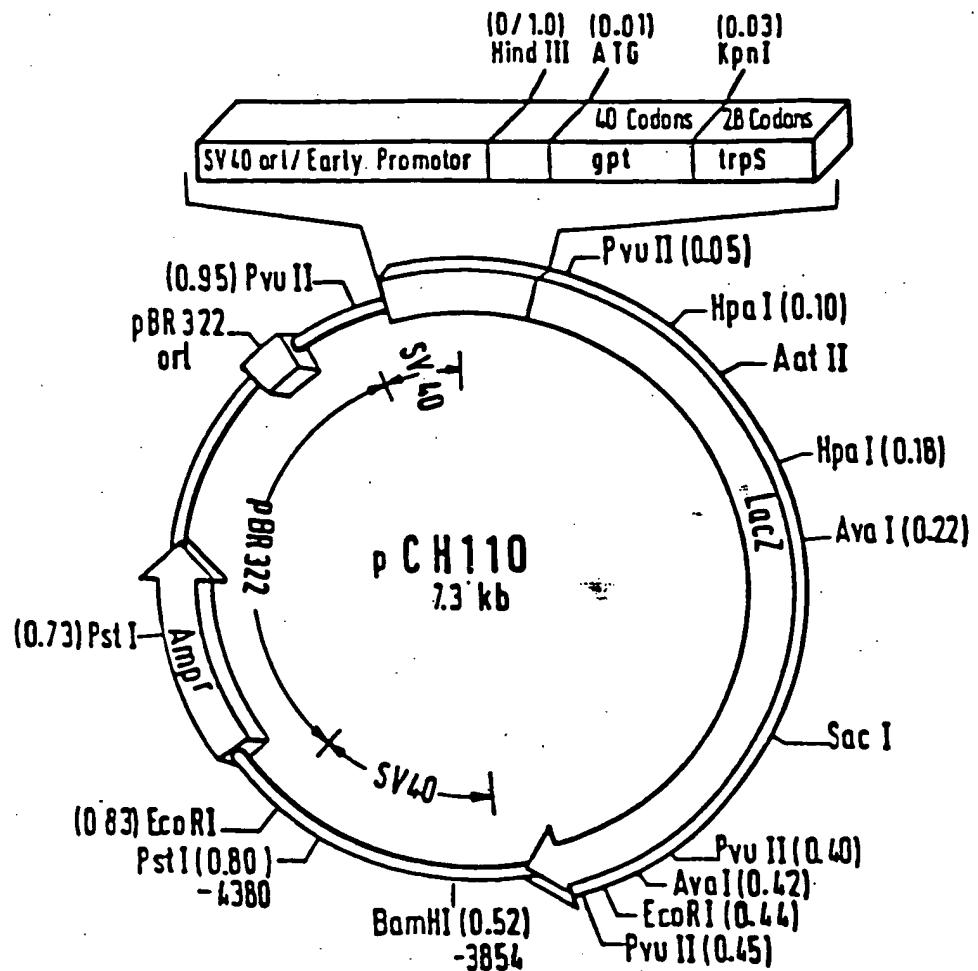
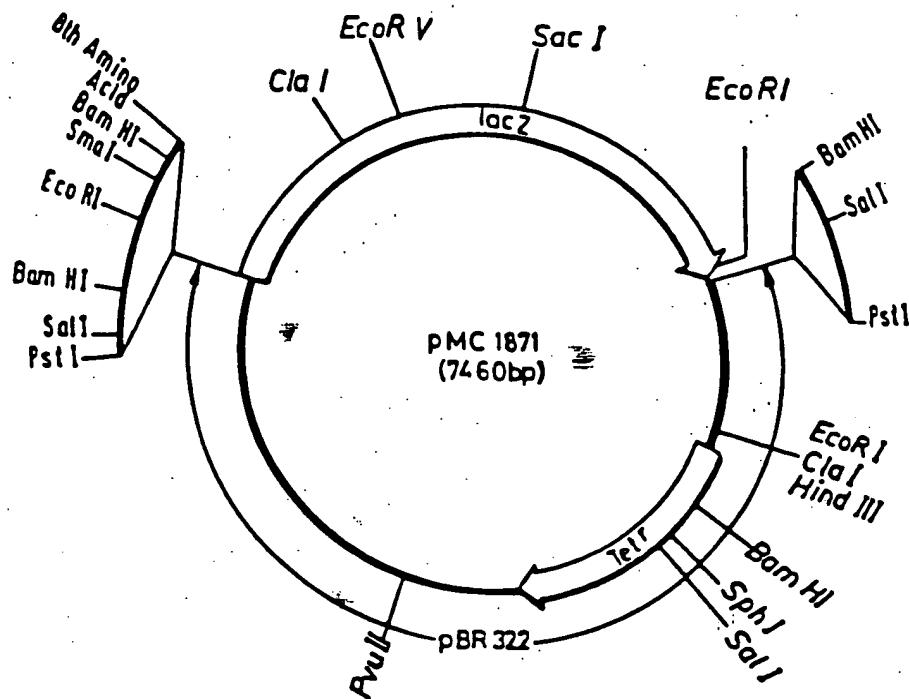


FIG. 5

FEUILLE DE REMPLACEMENT

7/12



pBR322 — PstI      SalI      BamHI      EcoRI      SmaI  
CT GCA GGT CGACGG ATC CGG GGA ATT CCC  
BamHI pro val      lys stop stop stop  
GGG GAT CCC GTC — lac Z — AAA TAA TAA TAA CCG GGC  
  
 AGG GGG GAT CCG TCG ACC TGC AG — pBR322  
Bam HI      Sal I      Pst I

FIG.6

FEUILLE DE REMPLACEMENT

8/12

1 2 3 4 5 6 7 8 9 10 M 0



FIG. 7

FEUILLE DE REMplacement

9/12

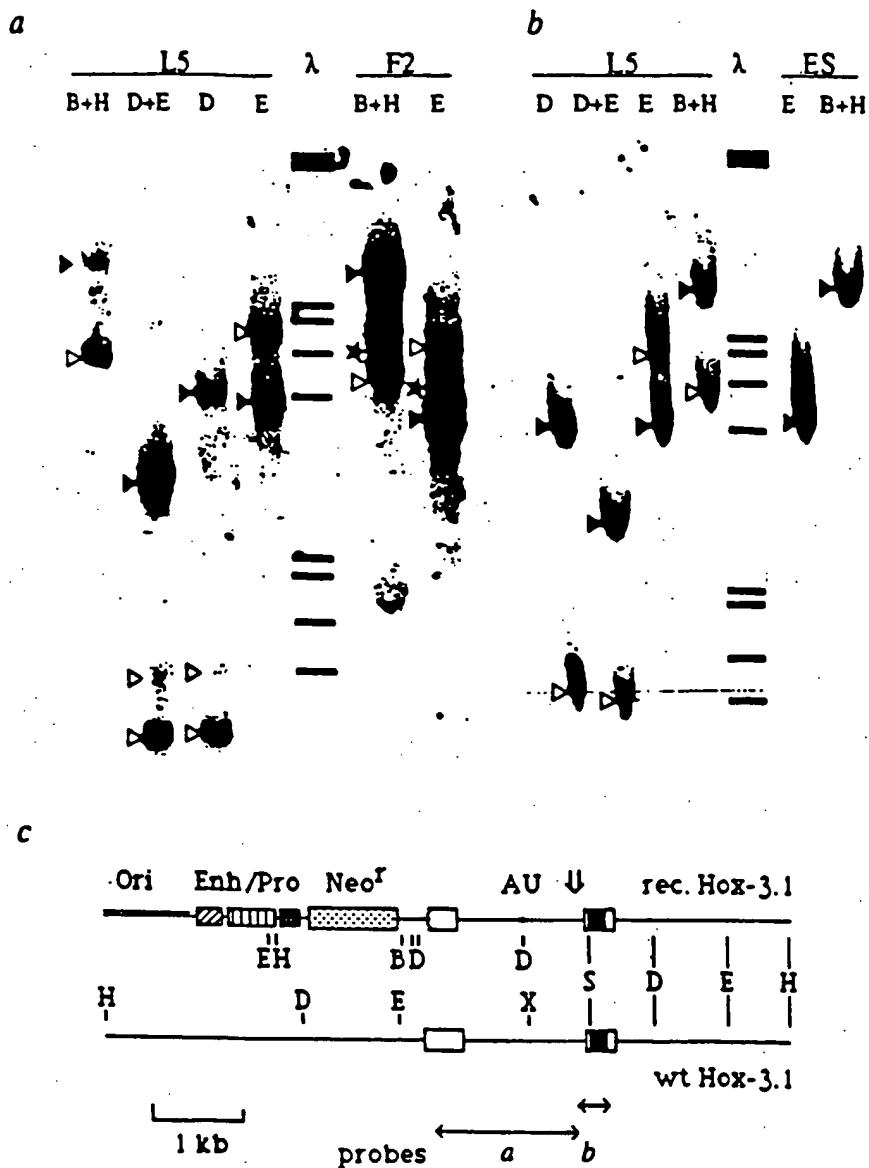


FIG. 8

FEUILLE DE REMplacement

10/12



Fig. 9A

**FEUILLE DE REMPLACEMENT**

11/12

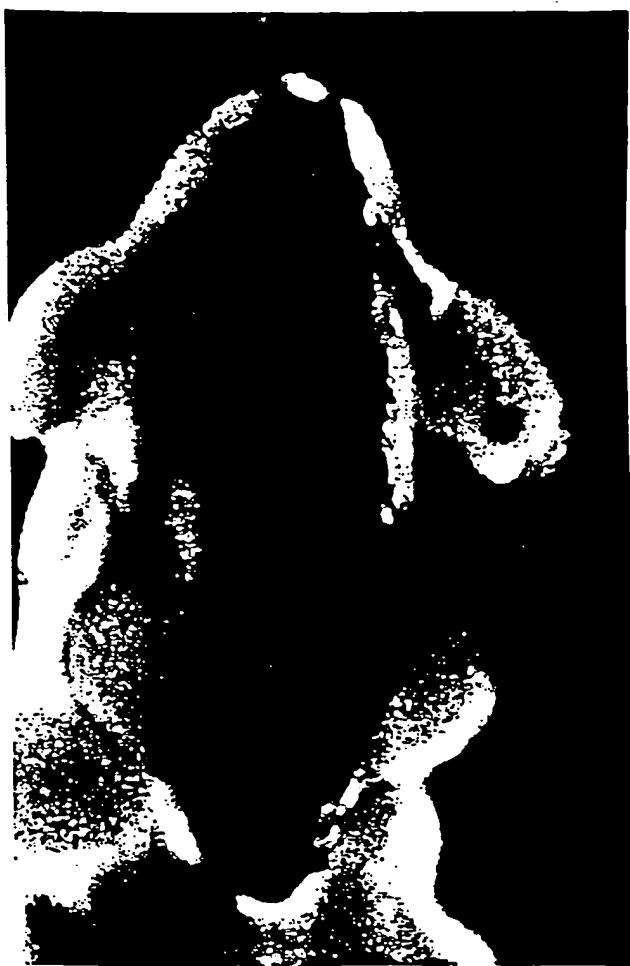


Fig. 9B

FEUILLE DE REMPLACEMENT

12/12



Fig. 9C

**FEUILLE DE REMPLACEMENT**